



Universitat de Barcelona  
Facultat de Biblioteconomia i Documentació

# La preservación de los recursos de hoy para los investigadores del futuro

**CSIC, Madrid, 26 de mayo de 2006**

**Alice Keefer [alice.keefer@ub.edu](mailto:alice.keefer@ub.edu)**

**Facultad de Biblioteconomía y  
Documentación**

**Universitat de Barcelona**

# **La preservación de los recursos de hoy...**

**0. Introducción**

**1. Principales retos**

**2. Contenido**

**3. Depósitos institucionales**

**4. Tendencias: preservación del material de investigación**

# La preservación de los recursos de hoy...

## 0. Introducción

# La preservación de los recursos de hoy...

## Introducción

Material de investigación - ¿qué contenido? :

- Producto final: *eprints*, artículos, informes, tesis, etc.
- Datos generados durante la investigación
- Información de posible interés para los investigadores futuros : es decir, TODO
- Información producida por instituciones de investigación: datos administrativos, etc.

# La preservación de los recursos de hoy...

## Introducción

### La preservación digital:

Es el conjunto de procesos destinado a garantizar el acceso y uso de los materiales durante el tiempo que se considere necesario.

### Incluye:

- la planificación
- la distribución de recursos humanos, económicos, institucionales
- la aplicación de los métodos y de las tecnologías considerados como las “mejores prácticas”

# La preservación de los recursos de hoy...

## 1 Principales retos:

1.1 técnicos

1.2 legales

1.3 económicos

1.4 institucionales

# La preservación de los recursos de hoy...

## Principales retos

### Principales amenazas tecnológicas:

- Obsolescencia tecnológica
- Fragilidad de los soportes
- Vulnerabilidad de la información digital

# La preservación de los recursos de hoy...

## Principales retos

Elementos informáticos que afectan a la legibilidad del recurso:

- Formatos de datos
- Programas de aplicación
- Sistema operativo
- Ordenador
- Soportes (ópticos y magnéticos)



# La preservación de los recursos de hoy...

## Principales retos

Ejemplos de la expectativa de vida de los soportes  
(Jones/Beagrie):

Soporte	25 HR 10°	30 HR 15°	40 HR 20°	50 HR 25°	50 HR 28°
D3 cinta mag.	50 años	25 años	15 años	3 años	1 año
DLT cinta mag. estuche	75 años	40 años	15 años	3 años	1 año
CD/DVD	75 años	40 años	20 años	10 años	2 años
CD-ROM	30 años	15 años	3 años	9 m.	3 m.

# La preservación de los recursos de hoy...

## Principales retos

### Vulnerabilidad de la información digital

Facilidad de pérdida:

- errores humanos
- malfuncionamiento de sistemas: equipos informáticos, redes de telecomunicaciones, electricidad, etc.
- acciones malintencionadas

La forma de combatir esta pérdida de información es mediante sistemas de seguridad, controles de acceso, copias de seguridad, etc.

# La preservación de los recursos de hoy...

## Principales retos

### Requisitos de la Cornell University:

Almacenamiento y mantenimiento de material digital antes de archivarse en el depósito institucional:

- Soportes
- Copias de seguridad
- Equipos de lectura y de grabación
- Almacenamiento
- Vigilancia y seguridad
- *Refreshing*
- Documentación

[www.library.cornell.edu/imls/image%20deposit%20guidelines.pdf](http://www.library.cornell.edu/imls/image%20deposit%20guidelines.pdf)

# La preservación de los recursos de hoy...

## Principales retos

### Aspectos legales:

- derecho a recopilar y archivar los recursos
- derecho a poner a disposición del público el material
- derecho a modificar los recursos archivados para garantizar su preservación.

# La preservación de los recursos de hoy...

## Principales retos

### Aspectos legales 2:

La responsabilidad de la biblioteca si el material contiene información:

- difamatoria
- confidencial o que no respeta la privacidad

# La preservación de los recursos de hoy...

## Principales retos

### Aspectos económicos:

- recursos tecnológicos
- recursos humanos
- recursos administrativos

.....¡a lo largo de los años!

# La preservación de los recursos de hoy...

## Principales retos

	Más fácil y ----->Más difícil y más económico más caro
Formatos	Pocos -----> Muchos
Derechos	Titularidad ----->Titularidad propia ajena
Control	Alto -----> Bajo

Granger, S., Russell, K. Weinberger, E.

# La preservación de los recursos de hoy...

## Principales retos

Políticas/actitudes *pro-activas*:

“la demora en la toma de decisiones hará que los requisitos de preservación sean más complejos, más laboriosos y, por ello, más caros.”

*Cedars guide to digital collection management*



# La preservación de los recursos de hoy...

## Principales retos

Etapas en el proceso de creación de un depósito institucional (*repository*):

- 1) reconocer el problema
- 2) actuar
- 3) consolidar
- 4) institucionalizar**
- 5) externalizar

(Cornell University)

# La preservación de los recursos de hoy...

## 2. Contenido

2.1 formatos

2.2 propiedades significativas

2.3 selección

# La preservación de los recursos de hoy...

## Contenido

### Text files

- Microsoft Word
- PDF

### Other Microsoft Office files

- Microsoft PowerPoint
- Microsoft Excel
- Microsoft Access

### Video/Animation files

- Flash
- QuickTime

*Table 5: File formats used by  
Academics at Loughborough  
University*

### Web based files

- HTML

### Image files

- GIF
- JPEG
- BMP

### Other files

- CAD
- FileMaker Database
- Hot Potatoes
- LaTeX

[Bates, Melanie et al. \*Digital lifecycles and file types: final report\*. JISC; Loughborough University, 2006.](#)

# La preservación de los recursos de hoy...

## Contenido

¿Qué material?

¿Qué atributos se quieren preservar?

- Datos
- *Look and feel*

¿Cómo se puede garantizar la integridad y la autenticidad?

Es decir,

- asegurar que el documento es lo que pretende ser
- asegurar que sus datos no han sido cambiados

# La preservación de los recursos de hoy...

## Contenido

Criterios tradicionales: valor, uso, pertinencia.

Nuevos criterios:

- que la institución tenga pleno derecho a manipular los datos para asegurar su acceso en entornos informáticos del futuro;
- que el recurso sea de un formato legible actualmente y previsiblemente en el futuro;
- que si el recurso esté en un soporte gestionable para su transferencia y/o almacenamiento;
- que el recurso disponga de documentación, incluyendo los metadatos.

# La preservación de los recursos de hoy...

## 3.Depósitos institucionales

# La preservación de los recursos de hoy...

## Depósitos institucionales

Variedad de nombres:

- *Institutional repositories* (ex., [Scholars' Bank](#), [Recercat](#))
- *Digital archives / repositories* (ex., [U.K.Data Archive](#))
- *E-print archives* (ex., [arXiv](#))

Entre el 2003 y el 2004 hubo un crecimiento del 100 % de los depósitos institucionales (Swan/Brown).

# La preservación de los recursos de hoy...

## Depósitos institucionales

¿Para qué se crea un depósito institucional? (Estos objetivos no son necesariamente compatibles):

- gestión de los bienes intelectuales institucionales para su reutilización
- almacenamiento de los recursos digitales
- promoción de Open Access y contribución a los cambios del sistema de publicación científica
- disponibilidad de una herramienta de administración (estadísticas, *márketing*, etc.)
- **preservación a largo plazo de los bienes digitales**



# La preservación de los recursos de hoy...

## Depósitos institucionales

Diferencias entre depósitos institucionales:

- la clase de recurso admitido
- la configuración del depósito
- las personas autorizadas a depositar materiales
- los formatos aceptados
- el tipo de acceso
- la forma de introducir los recursos

# La preservación de los recursos de hoy...

## Depósitos institucionales

### *Digital Repositories Roadmap* (abril 2006)

<[http://www.jisc.ac.uk/uploaded\\_documents/rep-roadmap-v15.doc](http://www.jisc.ac.uk/uploaded_documents/rep-roadmap-v15.doc)>

*There is little consensus on the future role of repositories for preservation, reflecting wider debate in this area. In particular, there are different views as to how far institutions, as opposed to national services, will be responsible for preservation. Some people see long-term digital preservation as an added value service layered onto the network of repositories, provided either by the institution itself or external service providers. On the other hand, some regard the institution as having only a short term responsibility for the curation of research outcomes...*

# La preservación de los recursos de hoy...

## Depósitos institucionales

### *Digital Repositories Roadmap (abril 2006)*

*...until these outcomes are formally published or stored in national data centres. From this perspective, the institution's primary responsibility is to give scholars an opportunity to access new material before waiting for the publication process, and to access data that would not be otherwise made available. ...[The] authors suggest that by **2010** there will be a firmer basis for a national preservation strategy that makes clear who has responsibility for preserving different types of data and who has responsibility for providing open access to resources.*

# La preservación de los recursos de hoy...

## Depósitos institucionales

¿Qué compromiso asume el depósito para la preservación a largo plazo?:

- todo el material depositado
  - sólo el material original científico
  - sólo el material conforme a los criterios institucionales
  - sólo el material en formatos recomendados
  - sólo el material nuevo / material antiguo
- o
- no asume ningún compromiso

# La preservación de los recursos de hoy...

## Depósitos institucionales

*Interactive assessment: selection of digital materials for long-term retention*

<<http://www.dpconline.org/graphics/handbook/dec-tree-doc.html>>

# La preservación de los recursos de hoy...

## Depósitos institucionales

*Trusted digital repositories* (depósitos digitales fiables)

<<http://www.rlg.org/longterm/repositories.pdf>>

Informe marco que define las características y responsabilidades de un depósito digital, elaborado por Research Library Group (RLG) y OCLC en 2002.

*Audit checklist for the certification of trusted digital repositories.* RLG; NARA, 2005.

<<http://www.rlg.org/en/pdfs/rlgnara-repositorieschecklist.pdf>>

# La preservación de los recursos de hoy...

## **4. Tendencias: preservación de materiales de investigación**

- 4.1 los resultados de la investigación
- 4.2 los recursos generados durante la investigación
- 4.3 otros materiales que utilizarán los investigadores

# La preservación de los recursos de hoy...

## Tendencias

LOCKSS (*Lots of copies keep stuff safe*) 1999-  
<<http://lockss.stanford.edu/>>

- sistema distribuido – aleatorio, descentralizado, en una variedad de plataformas informáticas
- almacenamiento y gestión in situ de los ficheros recibidos por suscripción a revistas
- comprobación de la integridad de los ficheros y detección y corrección de errores mediante un software *peer-to-peer*



# La preservación de los recursos de hoy...

## Tendencias

### CLOCKSS 2006:

Iniciativa para asegurar el acceso a largo plazo a la información digital. Agentes implicados: editores, bibliotecas y sociedades científicas.

Durante 2 años comprobarán si la tecnología y el modelo LOCKSS servirán para:

- mantener un archivo “oscuro” (*dark archive*)
- desarrollar un sistema *aceptable* para acceder a obras “huérfanas” (*orphan works*).

# La preservación de los recursos de hoy...

## Tendencias

Portico 2005-

<<http://www.portico.org/>>

Misión: preservar la documentación científica publicada en formato digital para asegurar que los científicos, investigadores y estudiantes accedan a ella en el futuro.

# La preservación de los recursos de hoy...

## Tendencias

### *Portico:*

Permitirá que las bibliotecas accedan a las revistas suscritas en momentos críticos (*trigger events*) como:

- desaparición del editor
- discontinuidad de la revista
- desaparición del archivo histórico del editor
- incapacidad del editor o de su equipo informático de dar acceso al contenido

# La preservación de los recursos de hoy...

## Tendencias

Holanda: Koninklijke Bibliotheek (KB)

<<http://www.kb.nl/kb>>

Ha firmado acuerdos formales para archivar, gestionar y preservar el contenido de revistas de grandes editoriales internacionales:

La KB actuará como *trusted digital repository*, poniendo el contenido a disposición de las bibliotecas suscriptoras.

# La preservación de los recursos de hoy...

## Tendencias

### E-ciencia

*In the next decade e-Science data generated from sensors, satellites, high-performance computer simulations,...scientific images, etc., will dwarf all of the scientific data previously collected in the whole history of scientific exploration.*

Beagrie, N. *A continuing access and digital preservation strategy..*  
(2002)

*...currently, there are no easy answers as how best to guarantee that digital data will last into the future without active management to preserve the data.*

Bleakly, D.R. (2002)

# La preservación de los recursos de hoy...

## Tendencias

### UK Data Archive

<<http://www.esds.ac.uk/aandp/create/data.asp>>

Depósito de material científico para:

- \* preservar en un archivo permanente los resultados de la investigación
- \* facilitar el acceso a estos datos para:
  - la investigación
  - la docencia y el aprendizaje
  - la validación de resultados

# La preservación de los recursos de hoy...

## Tendencias

### Digital Academic Repositories (DARE)

<http://www.darenet.nl/en/page/language.view/home>

Iniciativa de las universidades holandesas para garantizar el acceso en el futuro a:

- *working papers*
  - *e-prints*
  - tesis
  - informes
  - **datos no elaborados**
  - ponencias
  - multimedia
- ... y metadatos

# La preservación de los recursos de hoy...

## Tendencias

¿Para qué preservar la web?

- Para ofrecer material valioso a los investigadores del futuro, de la misma manera que ahora disponen de los documentos históricos.
- Para evitar que los sitios web, que son muy vulnerables, desaparezcan.

***...the life of an average website [is] estimated to be around 44 days (about the same lifespan as a housefly)***

*UK Web Archiving Consortium*



# La preservación de los recursos de hoy...

## Tendencias

UK Web Archiving Consortium

[www.webarchive.org.uk](http://www.webarchive.org.uk)

- Proyecto sobre *harvesting* selectivo de la web de la UK en el que cada miembro elige las webs que archivará.
- Incluyen la selección de webs sobre determinados acontecimientos:
  - programados: elecciones de 2005
  - inesperados: [atentado en el metro de Londres, 2005](#); tsunami del 2004.

# La preservación de los recursos de hoy...

## Tendencias

### Library of Congress – [Minerva Project](#)

- Pretende garantizar el acceso a materia prima para el Congreso y los investigadores, actualmente y en el futuro.
- Está dirigido por un equipo multidisciplinario con conocimientos de catalogación, derecho, servicios públicos y tecnología.
- Realiza estudios sobre la evaluación, la selección, la recopilación, la catalogación y la preservación de los recursos.

# La preservación de los recursos de hoy...

## Tendencias

International Internet Preservation Consortium

[www.netpreserve.org](http://www.netpreserve.org)

Comenzado en 2003, está formado por las bibliotecas nacionales de Australia, Canadá, Dinamarca, Estados Unidos, Finlandia, Francia (líder), Islandia, Italia, Noruega, Reino Unido, Suecia y el Internet Archive

# La preservación de los recursos de hoy...

## Tendencias

International Internet Preservation Consortium  
[www.netpreserve.org](http://www.netpreserve.org)

Objetivos:

- Ofrecer un forum para compartir conocimiento sobre la forma de archivar el contenido en la web
- Desarrollar y recomendar normas
- Desarrollar herramientas y técnicas para adquirir, archivar y dar acceso a sitios web
- Difundir información sobre los diferentes aspectos de la preservación de la web mediante congresos, publicaciones, etc.

# La preservación de los recursos de hoy...

## 5. Conclusión

# La preservación de los recursos de hoy...

## Conclusión

Nuevas soluciones pueden derivar de:

- la evolución de la tecnología que ofrecerá más alternativas,
- el desarrollo y la aplicación de normas para guiar el camino,
- la coordinación entre todas las entidades implicadas (*stakeholders*)

# La preservación de los recursos de hoy...

## Conclusión

Unas recomendaciones para la creación de un depósito para la preservación de material digital a largo plazo:

- Proceder con pasos lentos y seguros
- Hacer un seguimiento de las experiencias de otros centros y aplicar las mejoras prácticas
- Hacer un seguimiento de la evolución del entorno informático para prever los cambios
- Conseguir el respaldo institucional a largo plazo
- Mantener una comunicación fluida con los investigadores

# La preservación de los recursos de hoy...

## Conclusión

Recordar: para que los investigadores colaboren en la creación de los depósitos institucionales, mediante el depósito de su trabajo, necesitarán:

- tiempo
- recursos
- conocimientos
- apoyo
- incentivos
- seguridad